

Dokumente dekonstruieren!

Ein Plädoyer für semantische Bibliotheken

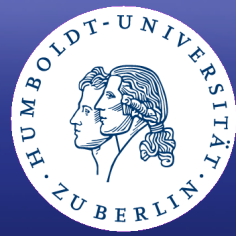
Prof. Dr. Stefan Gradmann
Humboldt-Universität zu Berlin / School of Library and Information Science
stefan.gradmann@ibi.hu-berlin.de

- Sind Bibliothekare Wissensarbeiter?
Traditionelle Bibliotheken und 'Dokumente'
- Zur De-Konstruktion des Dokumentbegriffs:
Von Xanadu zu OAI-ORE zu 'Semantic Publishing'
- Wie werden aus Bibliothekaren
Wissensarbeiter? Semantische 'Bibliotheken'
und was deren Betreiber wissen und können
müssen

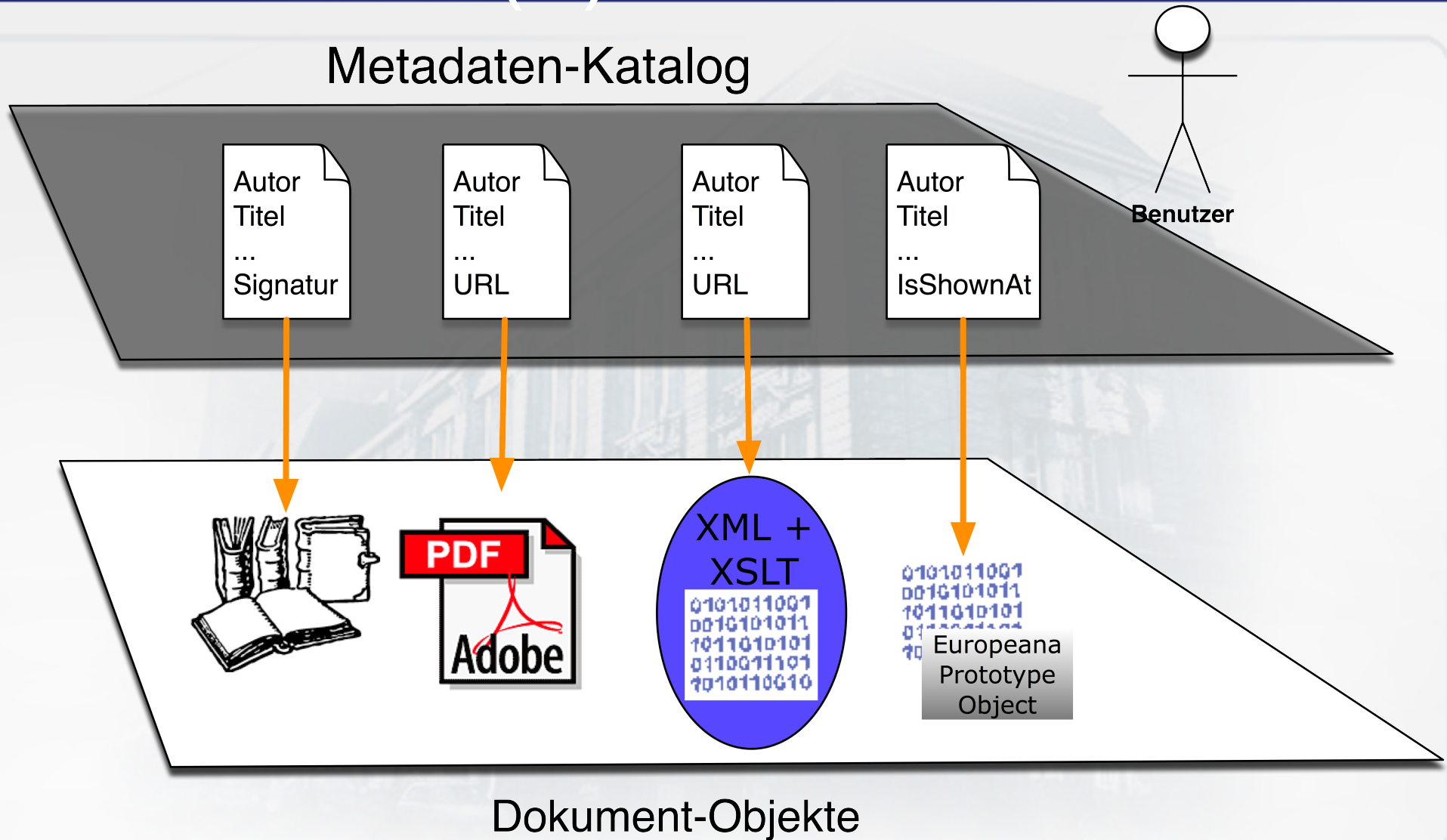
Sind Bibliothekare Wissensarbeiter?

Traditionelle Bibliotheken und 'Dokumente'

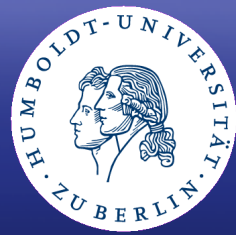
Funktionsprinzipien von Bibliotheken (1)



Metadaten-Katalog



Funktionsprinzipien von Bibliotheken (2)



- **Kataloge** vermitteln **Zugang** zu Informationsobjekten
- **Vermittelnde, unidirektionale 'Zeiger'** von Metadaten auf Objekte
- Objekte sind Teil eines '**Bibliotheksbestandes**'
- Die interne Bearbeitungslogik fokussiert auf
 - Objekte als Informations**container**,
 - Nicht so sehr auf den **Inhalt** dieser Container.
- **Einarbeitung, Speicherung, Beschreibung** und **Bereitstellung** von Informationscontainern sind die funktionalen Makro-Primitive

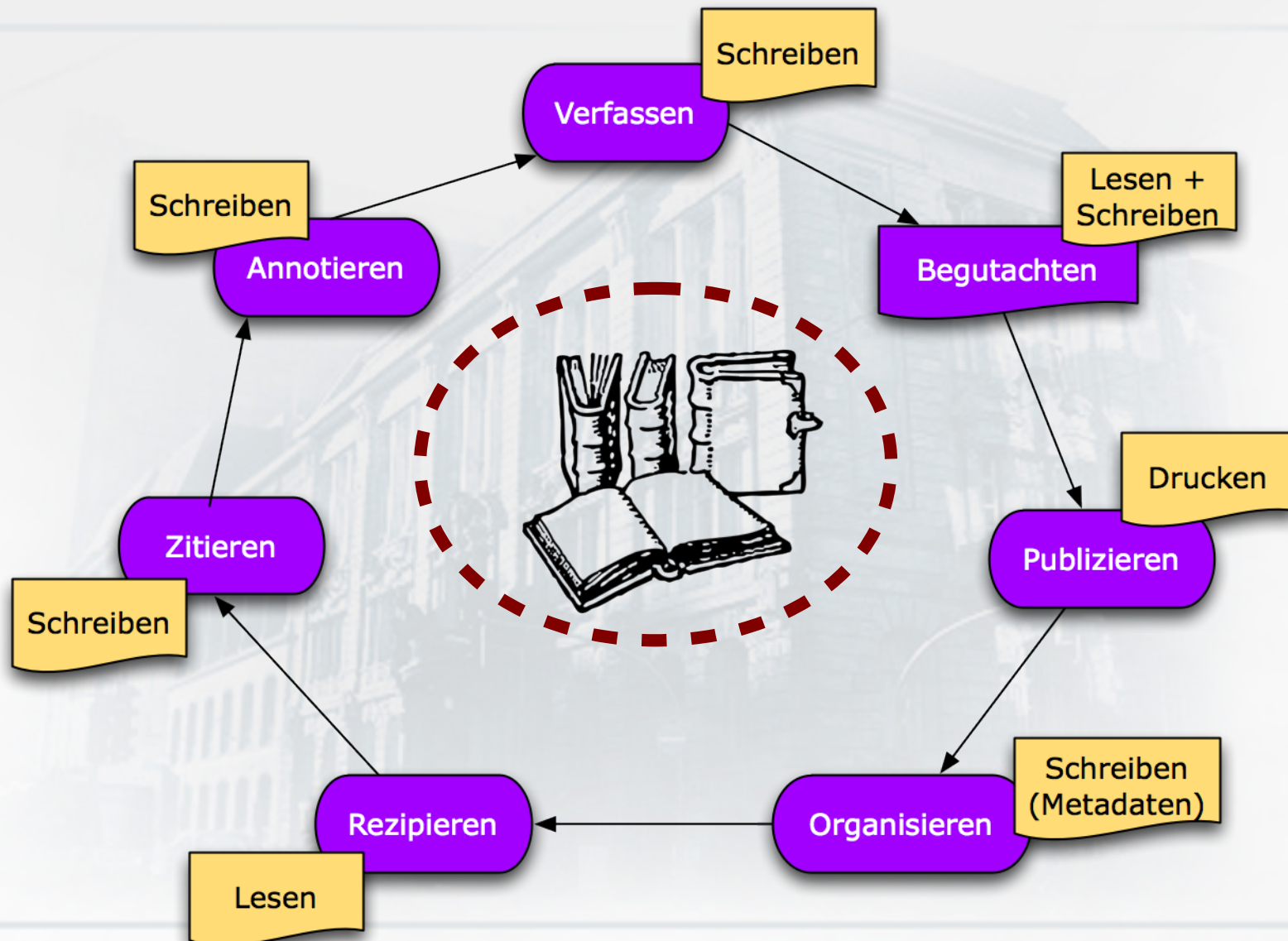
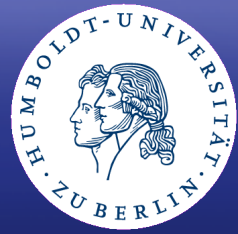
=> Bibliotheken waren so besehen traditionell eben **keine** Orte der Wissensproduktion (mit Ausnahmen ...)

=> Bibliothekare sind traditionell **keine** Wissensarbeiter!

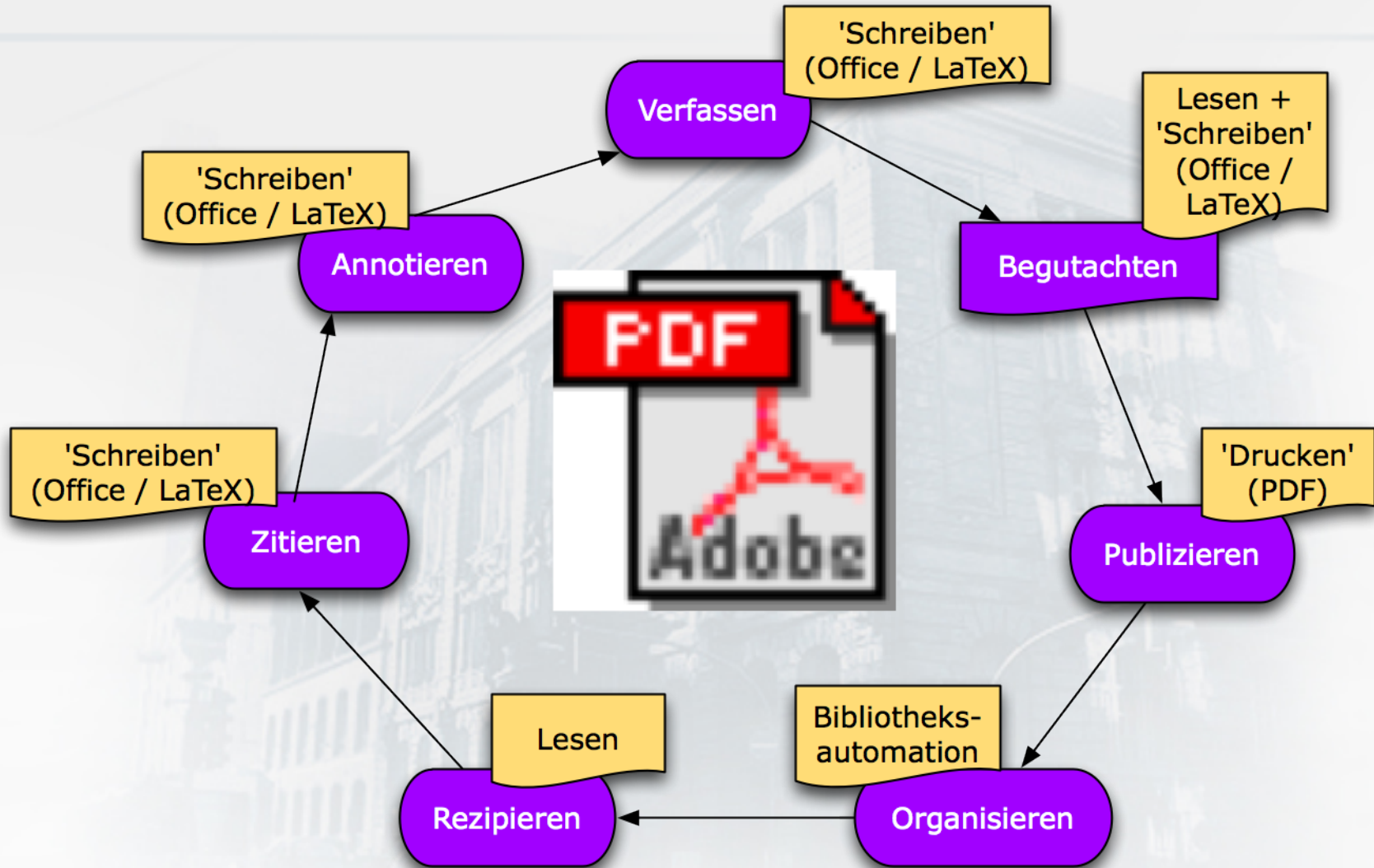


Zur De-Konstruktion des Dokumentbegriffs: Von Xanadu zu OAI-ORE zu 'Semantic Publishing'

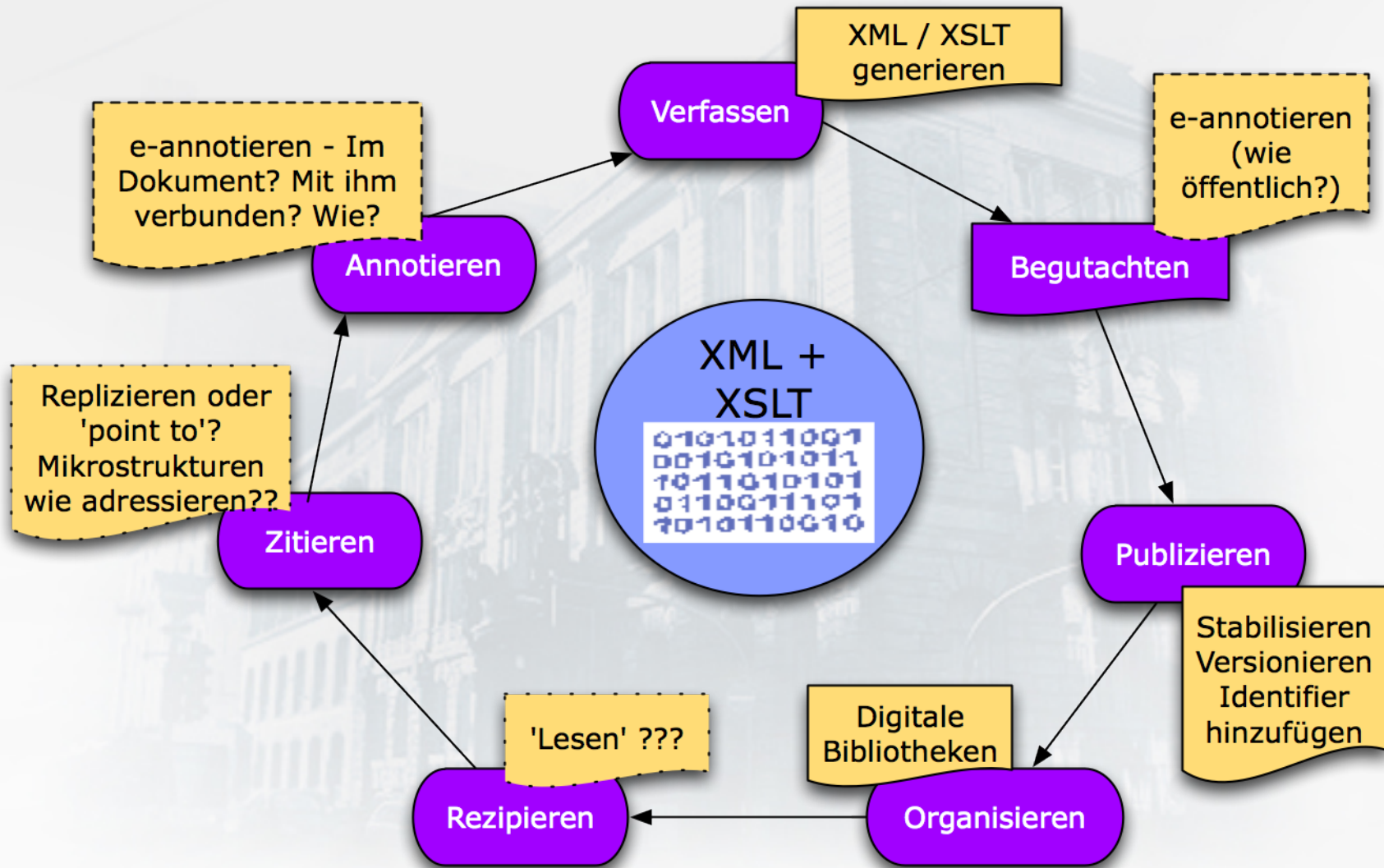
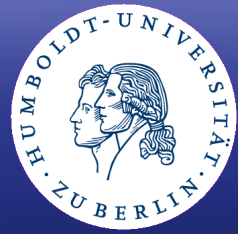
Das lineare Dokumentkontinuum in der Gutenberg-Galaxis



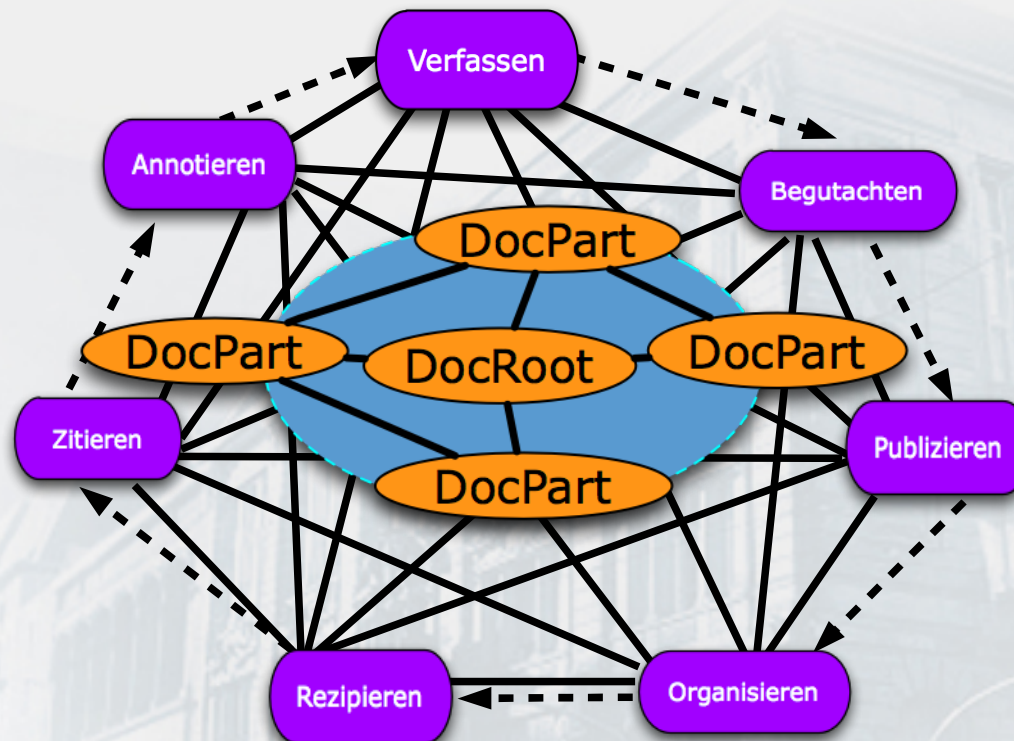
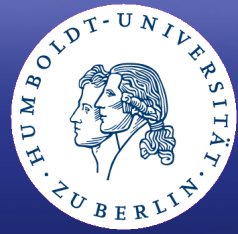
Das lineare Dokumentkontinuum im Emulationsmodus



Das lineare Dokumentkontinuum beim Eintritt in die Turing-Galaxis

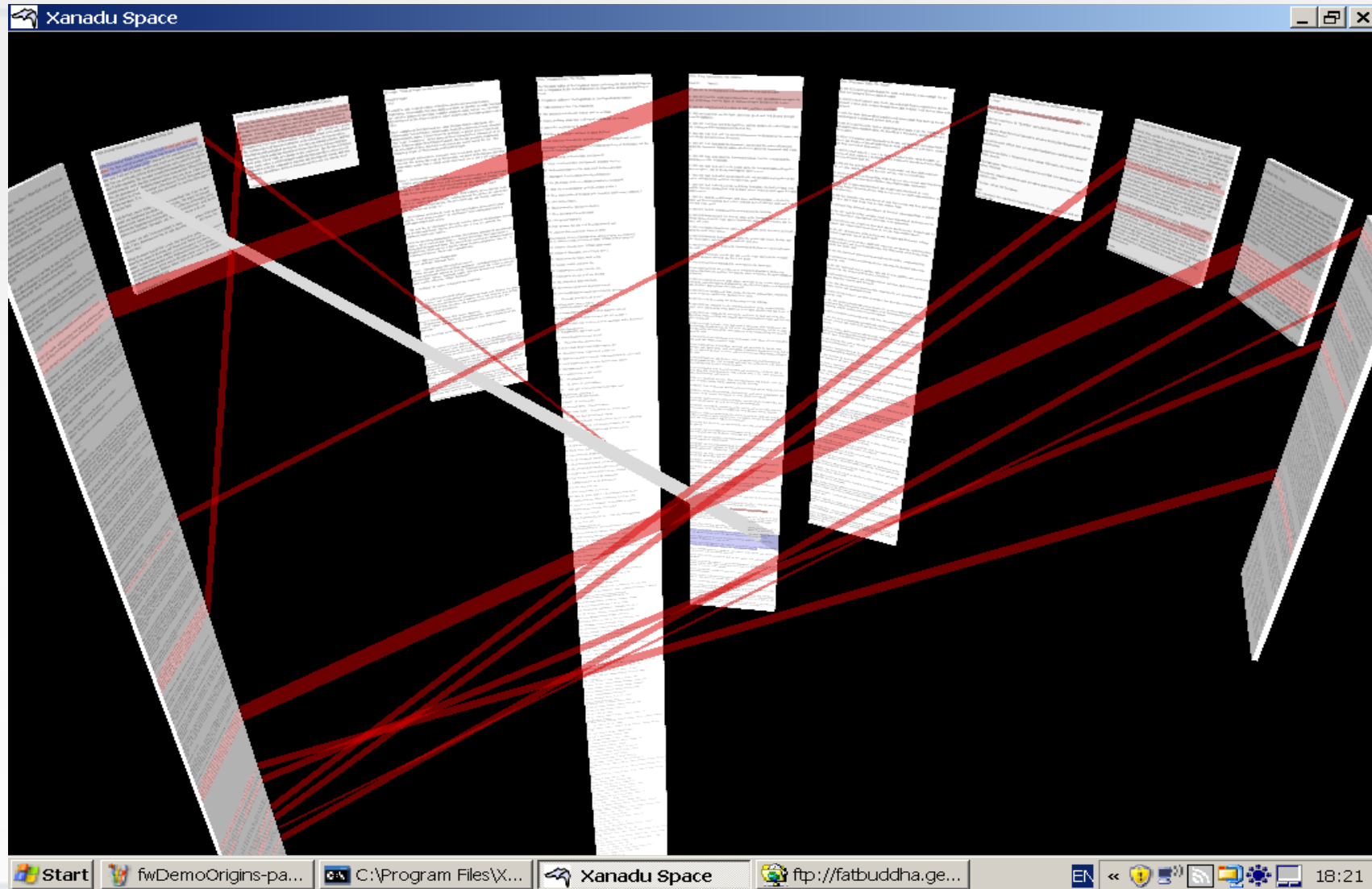
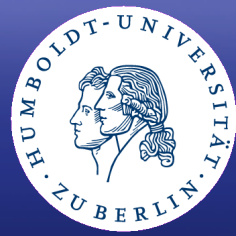


Ein WWW-basiertes Dokumentkontinuum als dreifacher Paradigmenwechsel

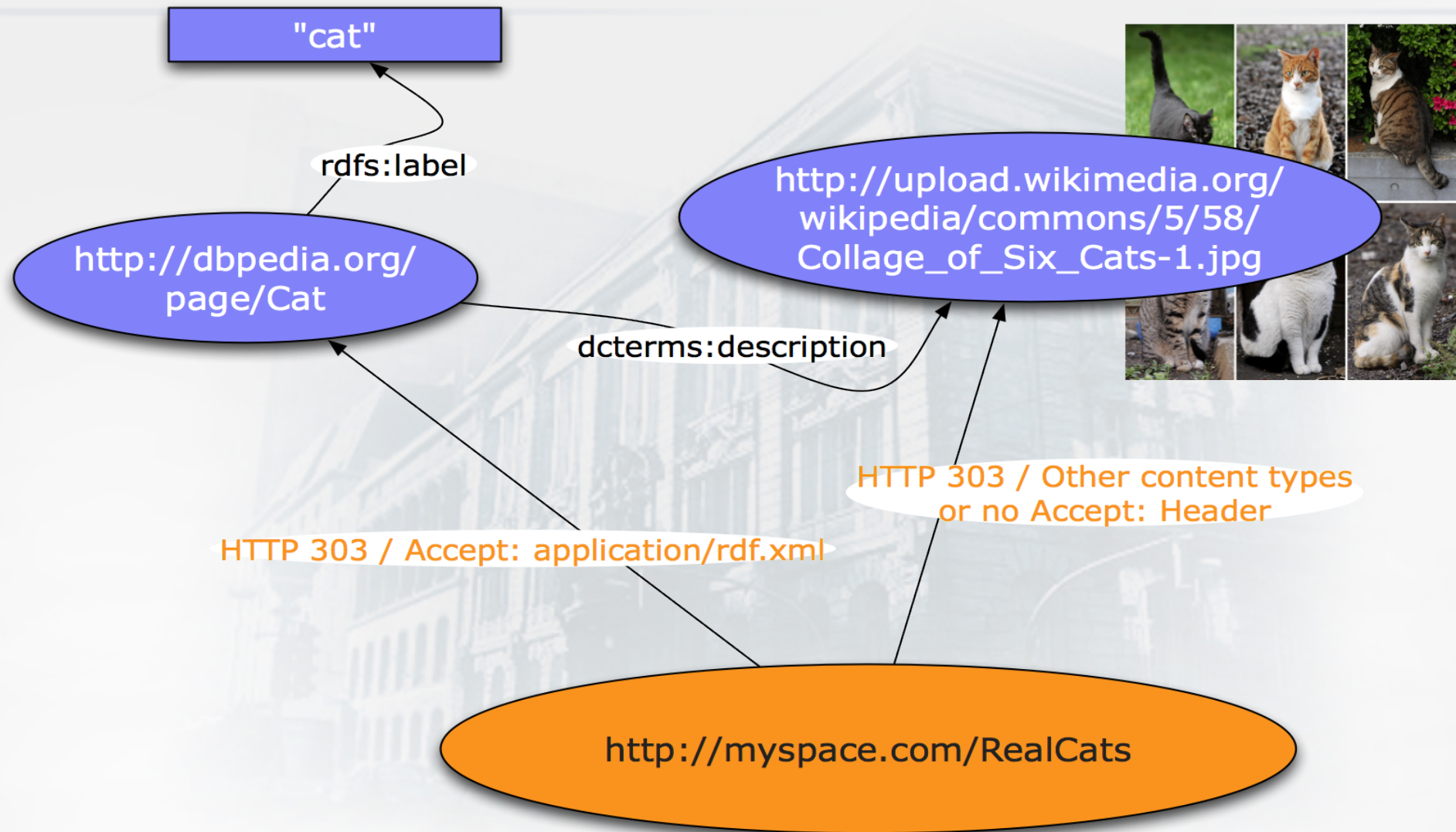
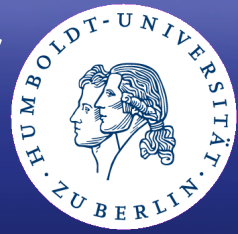


- Abnahme der funktionalen Determiniertheit durch traditionelle Kulturtechniken
- Auflösung des sequentiell-zirkulären Funktionsparadigmas
- Beginnende Erosion des monolithischen Dokumentbegriffs im Hypertext

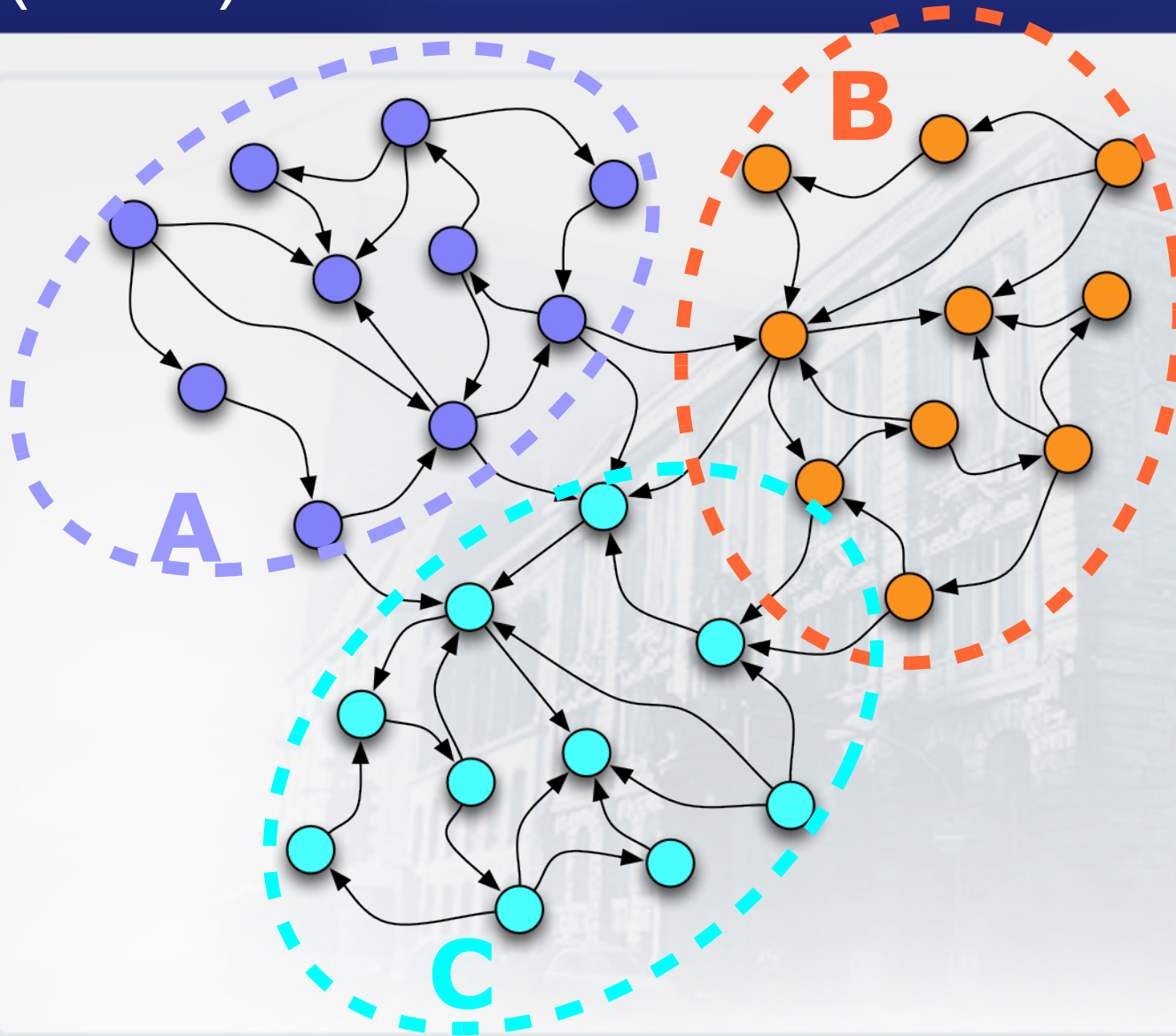
Xanadu (Ted Nelson): radikalisierter Hypertext ...



... und die Erweiterung des 'WWW der Dokumente' um ein Web der Dinge...



... führen zu 'Publikations'-Aggregationen (ORE) von 'Dokumenten' und 'Dingen'!

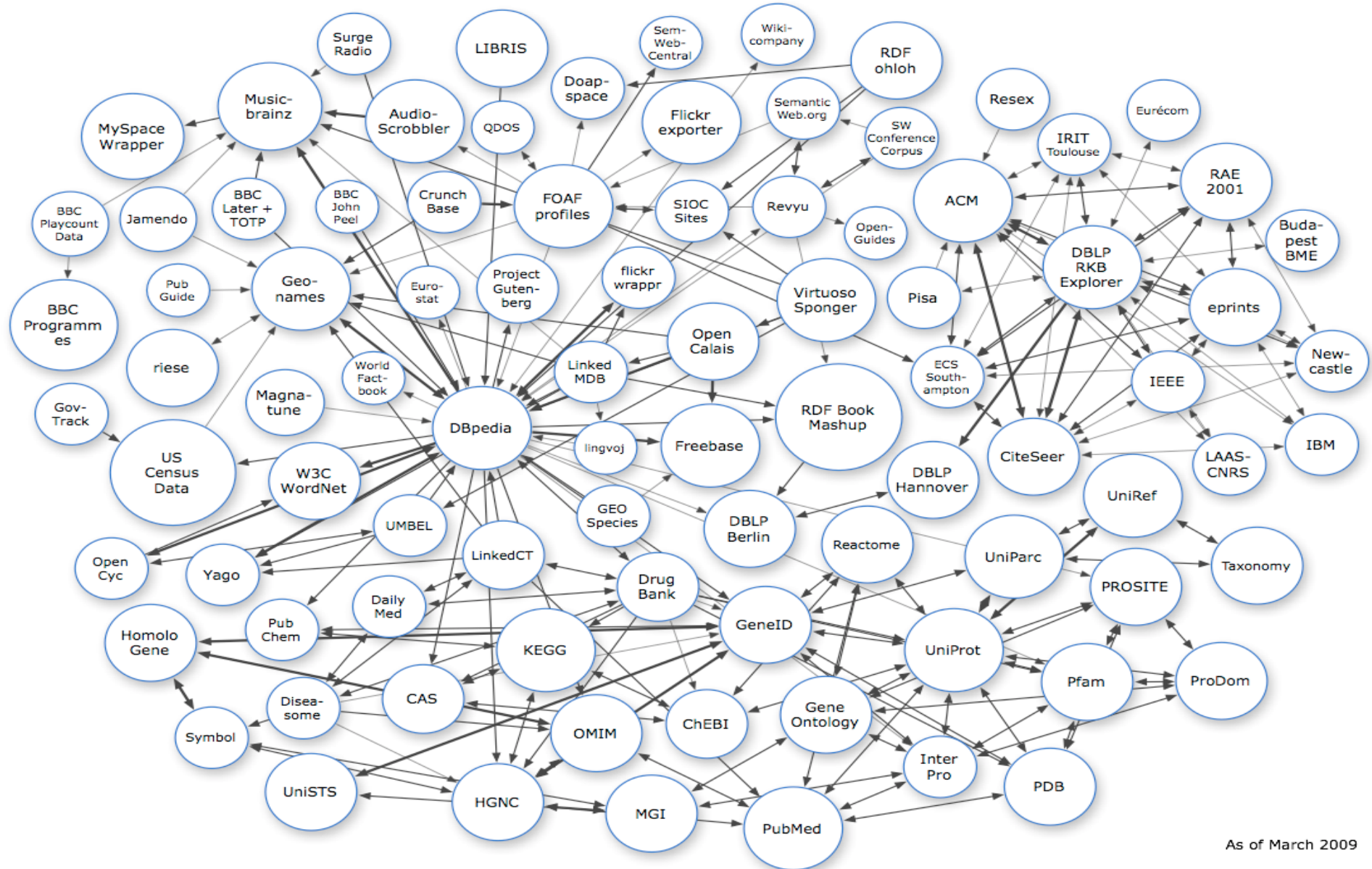


- Wo beginnen *resource aggregations*? Wo enden sie?
- Was konstituiert Dokumentgrenzen??
- Und welcher Knoten war zum Zeitpunkt X mit welchem verknüpft???

Damit werden Publikations-Aggregationen Teil des LoD



<http://esw.w3.org/TaskForces/CommunityProjects/LinkingOpenData/DataSets>



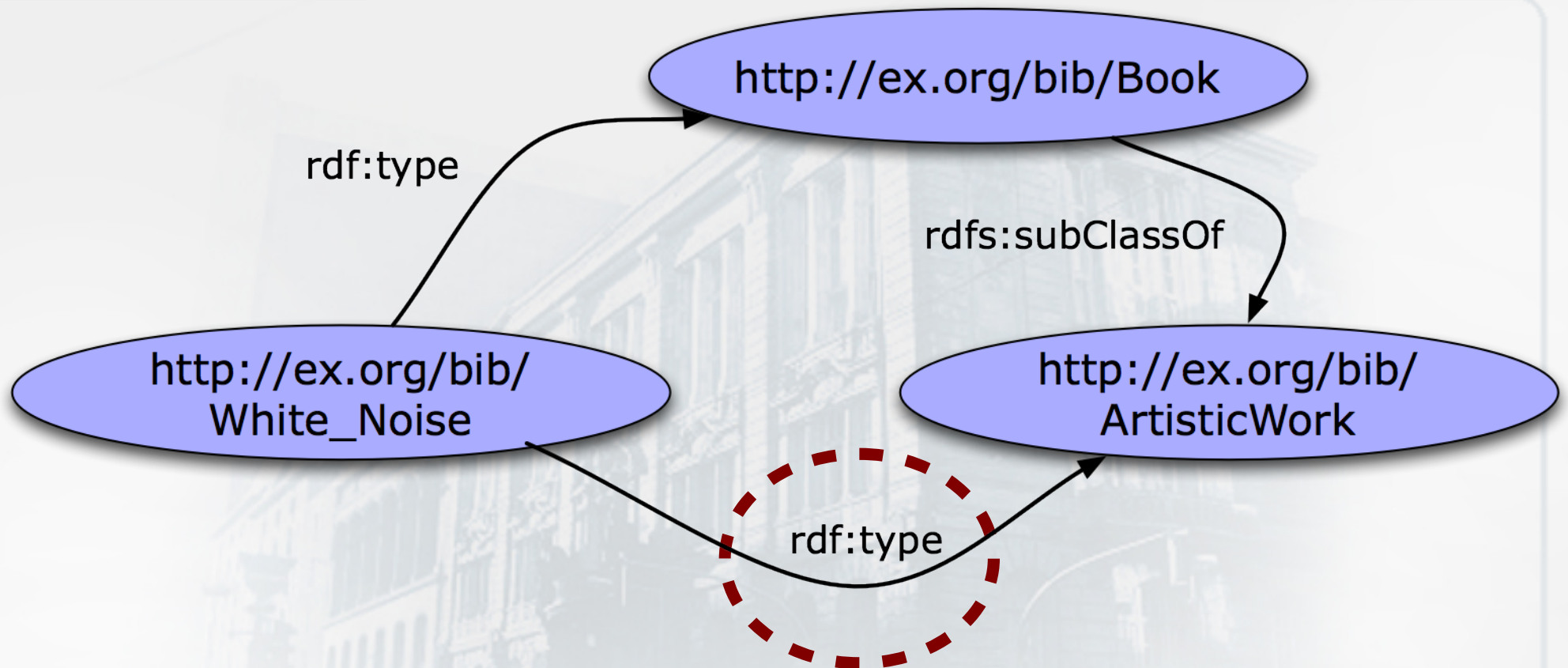
Tripel-Mengen und logische Schlüsse (1)



Tripel-Mengen und logische Schlüsse (2)



Tripel-Mengen und logische Schlüsse (3)



Dieser Typ Operationen eröffnet neues Potential für digitale Heuristiken

Vom Nutzen der Inferenz

Citation: van Haagen HHHBM, 't Hoen PAC, Botelho Bovo A, de Morrée A, van Mulligen EM, et al. (2009) Novel Protein-Protein Interactions Inferred from Literature Context. PLoS ONE 4(11): e7894. doi:10.1371/journal.pone.0007894 / Example provided by Jan Velterop



The screenshot shows the PLoS ONE article page for the paper "Novel Protein-Protein Interactions Inferred from Literature Context" by van Haagen et al. (2009). The page is highlighted with a red border. The title "Novel Protein-Protein Interactions Inferred from Literature Context" is circled in red. The authors listed are Herman H. B. M. van Haagen^{1*}, Peter A. C. 't Hoen¹, Alessandro Botelho Bovo², Antoine de Morrée¹, Erik M. van Mulligen¹, Christine Chichester¹, Jan A. Kors¹, Johan T. den Dunnen¹, Gert-Jan B. van Ommen¹, Silvére M. van der Maarel¹, Vinícius Medina Kern², Barend Mons¹, and Martijn J. Schuemie¹. The abstract states: "We have developed a method that predicts Protein-Protein Interactions (PPIs) based on the similarity of the context in which proteins appear in literature. This method outperforms previously developed PPI prediction algorithms that rely on the conjunction of two protein names in MEDLINE abstracts. We show". The page also features a search bar, navigation links, and various interactive elements like "Download", "Metrics", and "Share this Article".

'Semantic Publishing'

- Die Unterscheidung Daten vs. Publikation wird zunehmend funktionslos
- Zumindest in den empirisch geprägten Wissenschaften lösen sich die heute maßgeblichen Containerformate auf ('Liquid Books')

=> Jan Velterop's Konzept der “Nano-Publications” oder Bill Town's Beispiele aus der Chemie (z. B. OreChem) und

- David Shotton (2009): Semantic Publishing. The coming revolution in scientific journal publishing. Learned Publishing Volume 22, No 2, 85–94, April 2009; doi:10.1087/2009202
- Barend Mons, Jan Velterop: Nano-Publication in the e-science era (<http://www.surffoundation.nl/SiteCollectionDocuments/Nano-Publication%20-%20Mons%20-%20Velterop.pdf>)
- Alan Renear, Carol Palmer (2009): Strategic Reading, Ontologies and the Future of scientific Publishing. In: Science, August 2009, p. 828 – 832.

Wie werden aus Bibliothekaren Wissensarbeiter?

Semantische 'Bibliotheken' und was deren Betreiber wissen und können müssen

Bibliothek – und was nun?



- Mit der Containerbindung der Inhalte verschwindet auch die **βιβλιοθήκη** wie wir sie kennen.
- '**Bestand**' und '**Katalog**' werden zunehmend leere Metaphern.
- Diese Entwicklung ist eine **Bedrohung** für container- und bestandsbezogen konstituierte Institutionen.
- Sie ist eine **Chance** für diejenigen, die sich ernsthaft auf eine Beschäftigung mit den Inhalten einlassen.
- Bibliothekare haben dem Semantic Web viel zu bieten:
 - Strukturierte, gepflegte **Ressourcen für die Wissensorganisation** (VIAF, LCSH, Rameau, SWD etc.)
 - **Kompetenz und Disziplin** in der Arbeit mit Wissensorganisationssystemen (KOS)
- Unser Institut möchte helfen, dies Potential zu heben!

Was muß sich ändern?



- Ein erneuter dreifacher Paradigmenwechsel
 - **Container** im Focus → **Inhalte** im Fokus
 - Metadaten-**Katatalog** → **semantisches Netz** kontextualisierter Aggregationen
 - **Speicherung** und **Bereitstellung** von Informationsobjekten → **Wissensgenerierung** durch semantische Extraktions- und Aggregationsdienste
- Dazu braucht man unter anderem
 - Vertrautheit mit einschlägigen **WWW-Techniken** (RDF, SKOS, CIDOC-CRM etc.)
 - Offenheit für generische, nicht selbstgemachte **Standards**
 - Rechtzeitiger **Abschied von ausgehöhlten Metaphern**
- Eine Ahnung des Ergebnisses vermittelt
<http://www.europeana.eu/portal/thoughtlab.html>!

